

CIÊNCIA DE DADOS

MANUAL SOBRE A CRIAÇÃO DE PAINÉIS DE *BUSINESS INTELLIGENCE* - BI

Secretaria-Geral – SEG

Secretaria de Planejamento e Gestão Estratégica – SEPG

Serviço de Ciências de Dados – SERCID

Escritório de Projetos de Tecnologia da Informação - EPROJTI

O *Business Intelligence* (BI), também chamado de inteligência de negócios e inteligência empresarial, consiste no processo de coleta, organização, análise e disseminação de informações que auxiliam na tomada de decisões estratégicas e no planejamento das ações correspondentes.

O BI, além de embasar a tomada de decisão, pode ser utilizado para monitorar a concorrência, alimentar pesquisas de *marketing* e subsidiar demandas estratégicas da empresa que necessitem de mais amparo técnico e informações.

Esse processo transforma o grande volume de informações existentes sobre determinado assunto em inteligência corporativa acessada de maneira fácil, em ambiente *web*.

Arquitetura Tradicional

O *Data Warehouse* é um tipo de sistema de gerenciamento de dados projetado para ativar atividades de BI e para lhes fornecer suporte.

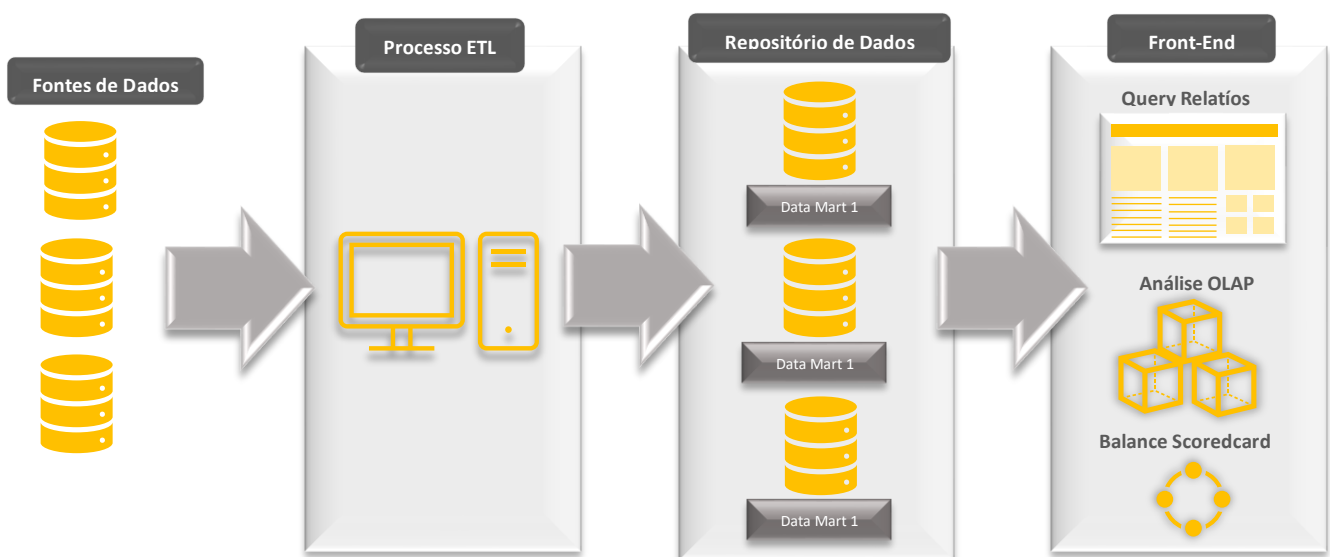
Destina-se exclusivamente a consultas e análises avançadas e geralmente contém grande quantidade de dados históricos derivados de fontes diversas, como arquivos de *log* de aplicativos e aplicativos de transações, os quais centraliza e consolida.

Os recursos analíticos dos *Data Warehouses* permitem que as organizações obtenham, em seus dados, informações de negócios úteis para melhorar a tomada de decisões.

Com o tempo, cria-se registro histórico – inestimável para cientistas de dados e analistas de negócios – que permite considerar um *Data Warehouse* como a “única fonte verdadeira” de dados sobre a organização.

O *Data Warehouse* típico normalmente inclui os seguintes elementos:

- banco de dados relacional, para armazenar e gerenciar dados;
- solução de extração, carregamento e transformação (ELT), para preparar os dados para análise;
- análise estatística, relatórios e recursos de mineração de dados;
- ferramentas de análise de clientes, para visualizar e apresentar dados aos usuários de negócios;
- outras aplicações analíticas mais sofisticadas, que geram informações úteis por meio de aplicação de algoritmos de *machine learning* e inteligência artificial (IA).



Os *Data Warehouses*, além de manterem registro histórico dos dados, oferecem o benefício único e abrangente de permitir que as organizações analisem grandes quantidades de dados variantes e deles extraiam valor significativo.

Isso só é possível, segundo o cientista da computação *William Inmon*, considerado o pai dos *Data Warehouses*, porque estes possuem quatro características únicas:

- São orientados para determinado assunto ou área funcional;
- São integrados; criam consistência entre diferentes tipos de dados de fontes distintas;
- Seus dados são estáveis ou não voláteis, ou seja, não mudam com o tempo;
- Analisam as variáveis ao longo do tempo.

Um *Data Warehouse* bem projetado realizará consultas muito rapidamente, fornecerá alta taxa de transferência de dados e flexibilidade suficiente para os usuários finais dividirem e organizarem, ou reduzirem, o volume de dados para exame mais detalhado, a fim de atender a várias demandas, seja em nível alto, seja em nível muito detalhado.

O *Data Warehouse*, finalmente, serve como base funcional para ambientes de BI de *middleware*, que fornecem aos usuários finais relatórios, painéis e outras interfaces.

O processo de construção de painéis de interação dinâmica é guiado, inicialmente, por reunião de alinhamento estratégico, na qual se intenta identificar os problemas e as necessidades da área demandante, a fim de que os principais indicadores possam ser priorizados.

Posteriormente, verifica-se a viabilidade dos respectivos indicadores em relação à existência de metadados estruturados nos sistemas. Existindo a estruturação básica, o ambiente analítico é concebido, e os respectivos painéis de interação dinâmica são elaborados.

Por fim, existe a necessidade de um processo de validação dos painéis pelas áreas de estatísticas. Em se tratando de painéis de dados processuais, são necessárias, ainda, as participações e as validações dos grupos gestores do 1º grau e do 2º grau.

Definição de escolha da publicação, caso sejam dados de gerenciamento interno ou externo.

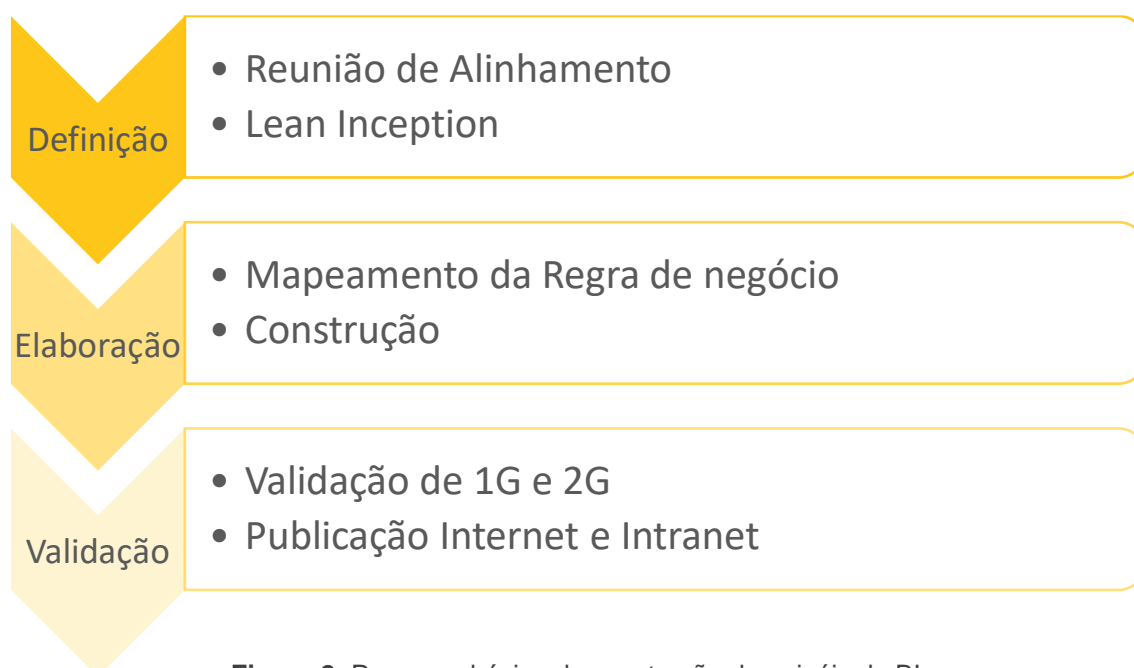


Figura 2: Processo básico de construção de painéis de BI

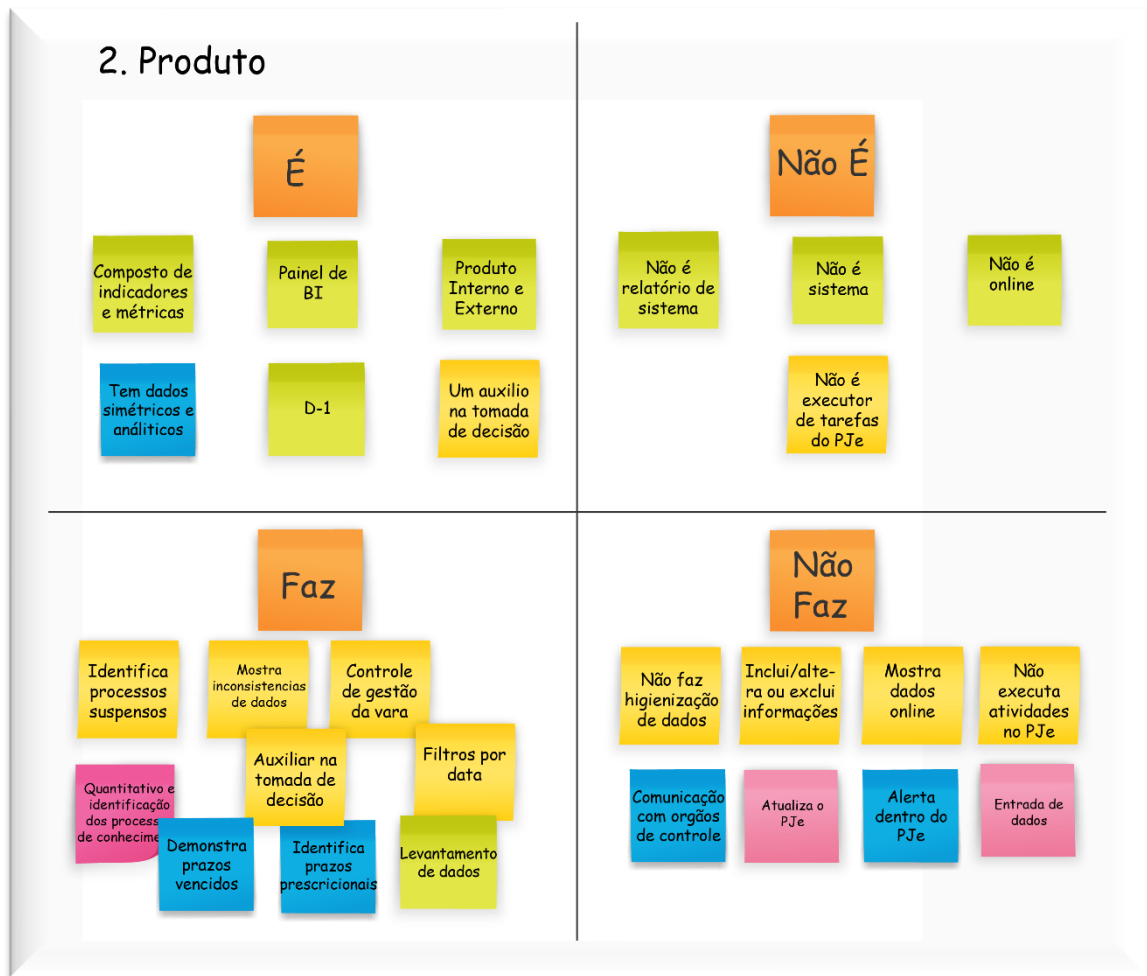


Figura 3: Definição do produto utilizando a técnica “É, Não É, Faz, Não Faz”

Ao tratar de BI, é quase obrigatório mencionar o processo de ETL – a sigla significa Extração, Transformação e Carga (em inglês, *Extract, Transform and Load*), que trabalha com toda a parte de extração de dados de fontes externas.

Essa transformação busca atender às necessidades de negócios e carga dos dados dentro do *Data Warehouse* ou *Data Mart* ou para demandas de importação e exportação de dados.

- **EXTRAÇÃO:** fase em que os dados são extraídos dos OLTPs e conduzidos para a *staging area* (área de transição ou área temporária), onde são convertidos para um único formato.
- **TRANSFORMAÇÃO:** é nesta etapa que são realizados os devidos ajustes para melhorar a qualidade dos dados e consolidar dados de duas ou mais fontes.
- **CARGA:** consiste em fisicamente estruturar e carregar os dados para dentro da camada de apresentação, seguindo o modelo dimensional.

Recomendações

- Carga dividida em níveis (Stage Area/RAW, Foundation/DW ou Data Mart e Application)
- As cargas oriundas do ambiente transacional (OLTP) devem ser realizada fora do horário de funcionamento do sistema, a fim de evitar possíveis sobrecargas;
- O ambiente de armazenamento OLAP deve ser segregado em relação ao storage, controladora, VLAN e instancia de banco de dados, a fim de evitar sobrecarga durante o processamento analítico;

- A instalação do banco de dados OLAP deve seguir as recomendações de blocos para ambiente analítico (32KB), entretanto depende do fabricante;
- A carga do primeiro nível (RAW/Stage Area) deve ser realizado de modo bulk, onde são realizadas cópias de blocos de dados de modo mais eficiente;
- A modelagem do ambiente DW deve inicialmente procurar o processo de modelagem estrela, caso não seja possível manter a modelagem pode-se derivar para o snowflake;
- A construção dos diferentes níveis de ETL deve seguir uma padronização de nomenclatura a ser definida pela organização;
- Evitar o uso de script's proprietários e linguagem de desenvolvimento dentro dos workflows;
- Observar as estratégias de inclusão e alteração em relação a necessidade de controle de histórico das dimensões, pois a mesma impacta em mais áreas de armazenamento.

Figure 1. Magic Quadrant for Data Integration Tools



Source: Gartner (August 2019)

Figura 4: Referência Gartner sobre ferramentas de ETL

Para que servem os Painéis de BI

Os painéis de *Business Intelligence*, um dos recursos mais populares dessas plataformas, são soluções de gerenciamento de informações e de visualização de dados usadas para analisar a instituição ou a empresa.

Neles, os dados são apresentados de modo mais fácil para a compreensão deles, e as informações que devam ser visualizadas são personalizadas. Os criadores de conteúdo podem usar elementos interativos, como filtros e ações, para combinar gráficos e relatórios em uma única tela, a fim de apresentar visão geral e rápida dos dados. Além disso, os resultados das análises podem ser compartilhados.

Os painéis, parte fundamental da estratégia de *BI* de uma organização, devem ser especialmente planejados e criados para permitir a análise dos conjuntos de dados relevantes, a fim de possibilitar a tomada de melhores decisões.

As plataformas de *BI* modernas permitem acessar, analisar, exibir e compartilhar dados com painéis hospedados na *web*, poupando aos analistas o trabalho de compilar planilhas manualmente.

Com uma ferramenta avançada e automatizada de *BI*, as partes interessadas podem criar painéis para fazer análises, tirar conclusões e tomar medidas. [tableau]

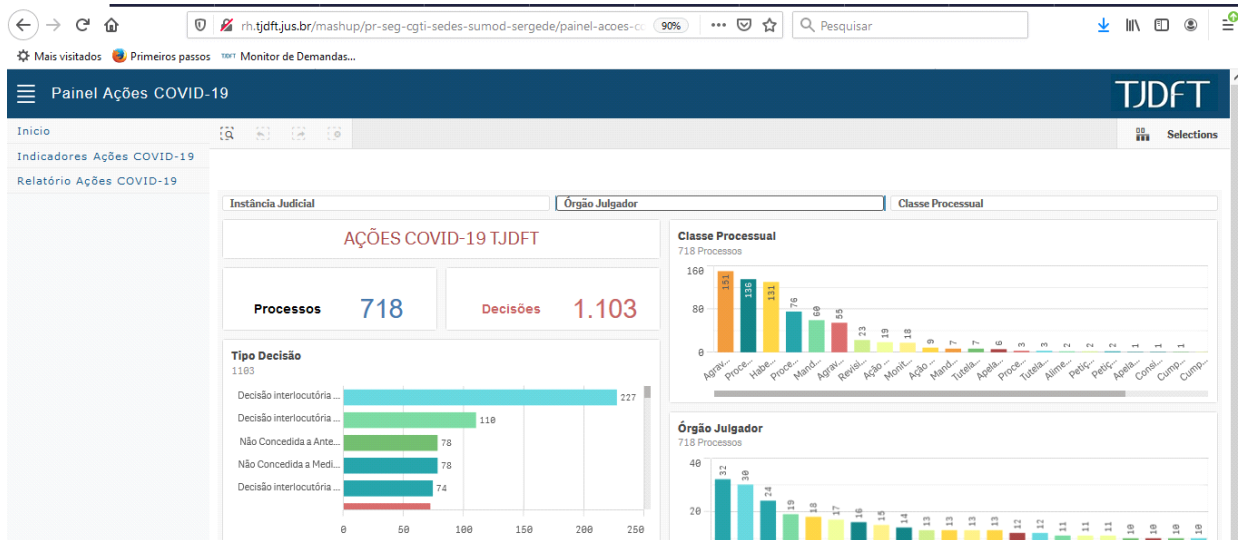


Figura 4: Painel Qlik Sence com dados sintéticos sobre COVID-19

Recomendações

- Dados sumarizados em indicadores estratégicos da área;
- Filtros customizados;
- Possibilidade de exportação do relatório analítico de dados;
- Importação das informações a partir do ambiente OLAP;
- Painel disponível em ambiente *web*;

Uma boa prática consiste na organização, em determinado ponto comum, dos painéis produzidos pela instituição, para que todos os dados estejam ali disponíveis e possam ser consultados por diferentes demandantes, bem como para que sejam evitadas possíveis redundâncias.



Estruturação da área de Ciências de Dados

Para consolidar as ações é importante estruturar uma área que concentre e trate as demandas, e que tenha algumas das seguintes atribuições:

- I – gerenciar a qualidade dos dados institucionais e proporcionar sua adequada extração;
- II - atender às demandas e aos projetos relativos a *Business Intelligence*, *Data Warehouse*, ETL e *Data Mining* da área judiciária;
- III - prestar apoio consultivo, inclusive sobre aspectos técnicos de coletas de dados e geração dos relatórios estatísticos;
- IV - prospectar ferramentas analíticas de *Business Intelligence*, *Data Discovery* e inteligência artificial;

- V** - elaborar painéis de informações estratégicas de apoio à tomada de decisão da Alta Administração;
- VI** – utilizar a inteligência artificial para realizar ações;
- VII** – apoiar a implantação da Lei Geral de Proteção de Dados no TJDFT.

Referências

[NewsMonitor] <https://www.newsmonitor.com.br/blog/business-intelligence-o-que-%C3%A9-e-qual-a-import%C3%A2ncia>

[Oracle] <https://www.oracle.com/br/database/what-is-a-data-warehouse/>

[VM2] <http://www.vm2.com.br/bi-business-intelligence?menu1=value>

[TABLEAU] <https://www.tableau.com/pt-br/learn/articles/business-intelligence/bi-dashboards>

[CETAX] <https://www.cetax.com.br/blog/etl-extract-transform-load/>